FAA-AM-71-28

AD 728844

A NON-VERBAL TECHNIQUE FOR THE ASSESSMENT OF GENERAL INTELLECTUAL ABILITY IN SELECTION OF AVIATION PERSONNEL

W. Dean Chiles, Ph.D.
Roger C. Smith, Ph.D.
FAA Civil Aeromedical Institute
P.O. Box 25082
Oklahoma City, Oklahoma 73125

June 1971

Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia 22151, for sale to the public.

Prepared for
DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
Office of Aviation Medicine
Washington, D.C. 20590

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FAA-AM-71-28		

4. Title and Subtitle
A NON-VERBAL TECHNIQUE FOR THE ASSESSMENT OF GENERAL INTELLECTUAL ABILITY IN SELECTION OF AVIATION PERSONNEL

5. Report Date
June 1971

6. Performing Organization Code

7. Author(s)
W. Dean Chiles, Ph.D.
Roger C. Smith, Ph.D.

8. Performing Organization Report No.

9. Performing Organization Name and Address
FAA Civil Aeromedical Institute
P. O. Box 25082
Oklahoma City, Oklahoma 73125

10. Work Unit No.

11. Contract or Grant No.

12. Sponsoring Agency Name and Address
Office of Aviation Medicine
Federal Aviation Administration
800 Independence Avenue, S. W.
Washington, D. C. 20590

13. Type of Report and Period Covered
OAM Report

14. Sponsoring Agency Code

15. Supplementary Notes
This research was performed under tasks AM-A-70-PSY-13, AM-A-71-PSY-13, AM-A-70-PSY-20, and AM-A-71-PSY-20.

16. Abstract
A study was conducted in which performance on a non-verbal problem-solving task was correlated with the Otis Quick Scoring Mental Ability Test and the Raven Progressive Matrices Test. The problem-solving task, called "code-lock" required the subject to determine the correct sequence in which to push five buttons in order to turn on a light. Measures of how quickly the subject responded and how many errors were made on each problem were taken from 45 college student volunteers. Results indicated substantial correlations (.50 to .60) between time measures on the code-lock task and the Otis but very limited relationships between the Raven and each code-lock measure. The implications of these findings for assessment of intellectual abilities are discussed.

17. Key Words
Intelligence
Problem-Solving
Selection
Personnel

18. Distribution Statement
Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia 22151, for sale to the public.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	8	$3.00

Form DOT F 1700.7 (8-69)

ABILITY CODES

SPECIAL

A

A NON-VERBAL TECHNIQUE FOR THE ASSESSMENT OF GENERAL INTELLECTUAL ABILITY IN SELECTION OF AVIATION PERSONNEL

I. Introduction.

Present selection procedures which test for ability to successfully complete training in many skilled aviation professions, such as air traffic control, give heavy weight to verbal skills. Performance on the selection task may therefore depend in large part upon the degree to which an individual has a history of language experiences appropriate to the verbal aspects of the assessment devices employed. Thus, selection procedures which include predominantly verbal components may discriminate against individuals with otherwise adequate abilities. Such discrimination is most likely to influence the performance of cultural minorities, since their language experiences may be markedly different from the experiences of most of the individuals used to standardize selection devices.[e.g., 7] Therefore, there is a need for a measure of job-relevant intellectual abilities which is not dependent on specific acquired verbal skills.

A device which may have some promise in this regard is the "code-lock" problem-solving task. The device was originally developed as part of a group problem-solving program for use in assessing air-crew performance,[2] and was subsequently modified for use with individuals. It requires a subject to determine the correct sequence for pressing push-buttons by using a specified search sequence. A modified version of the code-lock task, called COTRAN, was used by Alluisi and Coates[1] to test the relationship between this general type of procedure and various measures of intelligence and achievement. The COTRAN procedure is considerably more complex than the code-lock task, as it consists of three phases. Phase one is a standard code-lock problem, phase two is a second standard code-lock problem whose solution is a transform of the first problem, and phase three is a code-lock problem which is to be solved by application of the transformation rule rather than through the search sequence. Alluisi and Coates found that some COTRAN measures established by factor analytic procedures correlated up to .52 with the Raven Progressive Matrices Test (RPMT), a non-verbal intelligence assessment technique, and to .36 with the Quantitative and Total scores of the Scholastic Aptitude Test (SAT). However, the pattern of correlations in their study suggested that the standard code-lock portions of COTRAN (phases 1 and 2) might be as closely related to the RPMT as measures from the entire COTRAN task. One purpose of the present study was to examine this possibility by testing subjects on the code-lock procedure alone, and comparing their performance to their scores on the RPMT.

A second interest of this study was to determine to what extent performance on the code-lock task might be related to more general measures of intelligence than the RPMT. One of the limitations of the RPMT is that it has been found to have a limited relationship to performance in most academic or occupational settings.[3] If the code-lock task is also found to be related to general, as well as to non-verbal, intellectual ability, then its value for use in selection batteries would be much enhanced. In this case, the measure of intelligence chosen was the Otis Quick-Scoring Mental Ability Test since it is brief, easily administered, often used in scholastic placement, and is reasonably correlated with other tests of intelligence.[e.g., 5]

The third concern of this experiment was to study the effects of problem difficulty upon the relationship between code-lock and intelligence test performance. The measure of difficulty employed was the complexity of each code-lock problem, defined as the number of button presses required by the search sequence to determine the correct problem solution. In other words, a problem solution which matched the search sequence exactly would be of lowest difficulty, while

the more the correct sequence departed from the search sequence, the more difficult the problem.

II. Method.

Subjects. The subjects for this experiment were 45 paid male volunteers recruited from among the students of the University of Oklahoma part-time employment pool.

Performance Task Apparatus. The apparatus for the task consisted of a peripheral modification of the group problem-solving task.[2] The modification comprised the substitution of a small metal box on which five push-buttons were mounted in a row for the single button per each of five subject panels in the group version.

Three indicator lights, with green, amber, or red jewels, provided feedback to the subject. The green light was used to indicate that a given problem had been solved. The amber light indicated to the subject that his depression of one of the buttons had registered in the programming and scoring circuitry (failure of this light to illuminate was a signal to repress the button to register the response choice). The red light served two functions. Its initial illumination indicated to the subject that a problem was being presented and that he should immediately begin to search for the solution. Subsequently, the red light provided error information in the following manner: any time any of the buttons was pushed while the red light was on, the light would go out. If the button depressed was the correct first button in a given problem sequence, the light would remain out when the button was released; if the button was incorrect, the red light would come back on when the button was released. Thus, the initial step in solving a problem was to try the buttons one at a time in a left-to-right sequence to determine which button was first for solving the problem at hand. Having found that button, the subject again searched in a left-to-right sequence for the next correct button. Thus, if the next button pushed after finding the first correct button was the second in the sequence, the red light would stay out; if the next button was out of sequence for that problem, the red light would be re-illuminated, and the number-one button would again have to be pushed in order to continue with the search for the second button. Once each of the five buttons had been pushed in the correct sequence, the green light would be illuminated signifying that the problem had been solved. After a ten-second delay, the same problem would be presented a second time, thus requiring the subject to remember the sequence during the inter-problem interval. Thus, each problem was solved twice, once by searching for the answer and once by re-entering the already obtained answer. Subjects were required to make all responses with the index finger of the right hand.

The 24 problem sequences selected from the 52 possible sequences were presented in a randomized order so that the subject had no prior information as to the correct sequence for any given new problem.

The left-to-right search procedure specified to the subjects was chosen as the most efficient approach in terms of memory load during search. Therefore, subjects were instructed always to begin with the left-hand button and search from left to right; the left-hand button was to be considered to follow the extreme right-hand button in the search sequence. For example, assume the correct sequence for a given problem to be button depressions in the order 4, 5, 3, 1, 2. A subject who was proceeding according to instructions would emit the following responses ("E" designating an error and "R" a correct response): 1E, 2E, 3E, 4R, 5R, 1E, 4R, 5R, 2E, 4R, 5R, 3R, 1R, 2R (green light).

The subject's performance was recorded by means of an automatic scoring system with a punch tape output. The time at which each event occurred was recorded to the nearest 1/100 second with respect to the beginning of the experimental session (time zero). The button pushed, whether or not it was in the correct sequence, and whether it was associated with an initial solution or a re-entered solution was also recorded. The punch tapes were then read by a computer which complied and analyzed the data. The measures used were as follows:

1. Mean time per response (initial and re-entered solution combined). This measure simply determined the average number of seconds it took a subject to make each one of his responses during the entire course of the experiment. It provides an index of the speed with which each subject responded.

2. Mean surplus errors (initial solution). The surplus-errors measure was derived as follows:

for any given problem sequence, the exact number of errors could be predicted for a subject who followed the prescribed search procedure. Thus, this measure was the mean (per problem) of the actual number of errors minus the expected number of errors for the initial solution of the problems. For example, referring back to the previously presented sample problem sequence, the expected number of errors in solving that problem was five.

3. Proportion of errors (initial and re-entered solution combined). This represents the number of errors a subject made divided by the entire number of responses made to the problem. In other words, it is a measure of one aspect of problem-solving efficiency.

4. Mean time per response (re-entered solution). This value is calculated only for the responses during the confirming solution trial, and provides a measure of response speed after the various problems had been solved.

5. Mean solution time after first correct response (initial solution). In this case, the measure indicates how long it took a subject to determine the correct solution to a problem once he had found the button which started the correct pressing sequence.

The problem difficulty was controlled after the fact. One problem could be solved with only one error; one problem could be solved with nine errors, the maximum number of errors for any of the problems. For the difficulty analysis, the 24 problems were divided into three subsets with the following ranges of expected numbers of errors: 1 to 4, 5 to 6, and 7 to 9; the number of problems falling in each of the three categories was 8, 9, and 7 respectively.

Intelligence Measurement. The Otis Quick-Scoring Mental Ability Gamma Test, Form C, was administered using the standard instructions from the manual with a 30-minute time limit. The Raven Progressive Matrices Test (1962 Revision) was also given with the standard instructions and a 40-minute time limit. (Brief descriptions of the Otis and RPMT are presented in Appendix 1.) For 23 of the subjects, the Otis was given first, the code-lock performance second, and the RPMT last. For the other 22 subjects, the order was reversed.

Procedure. The subject, upon arrival at the laboratory, was given the paper-and-pencil test that had been chosen at random for him to take first. After finishing that test, or upon expiration of the time limit, the subject began the training period on the code-lock task. He was told to push only one button at a time using the index finger of his right hand. He was told to work as rapidly as he could but to avoid making unnecessary errors. The experimenter "talked the subject through" two complete problems and then stood behind the subject for a number of problems that was dependent upon when the subject had demonstrated that he understood the nature of the task and what was expected of him (this number was typically about three problems). The experimenter then entered the adjoining room from which he could observe the subject through a half-silvered mirror. The subject was permitted to work a total of at least ten practice problems; when required the number was extended to 11 or 12. When necessary during the training phase, the experimenter reminded the subject to follow the prescribed search sequence. No interaction with the subject was carried out during the test session itself. The duration of the test session depended on how fast the subject worked the 24 problems, both initial and second solutions; the mean for the 45 subjects was 25.37 minutes with a standard deviation of 2.10 minutes. (This mean included 7.83 minutes which was the sum of the fixed between-problem times.) The subject then took the remaining paper-and-pencil test.

III. Results.

Intelligence Measures. The obtained distribution of scores on the RPMT (Table 1) ranged from 14 to 35 with a mean of 24.9 and a standard deviation of 4.6. A score of 14 on this test represents the 75th percentile according to the published norms for age 20. The 90th percentile and 95th percentile are represented by scores of 21 and 24 respectively. In this study, 36 of the subjects (80%) scored at the 90th percentile or above. Thus, the range of RPMT scores was rather limited.

The scores on the Otis ranged from 43 to 76 with a mean of 61.6 and a standard deviation of 9.2 (Table 1). The mean for the Otis at age 18 or older is 42. Since a Gamma IQ of 100 is the equivalent of the 50th percentile, the range of obtained IQ scores (from 101 to 134) represents a variation from average to superior perform-

TABLE 1.—Frequency distribution of scores on the Raven Progressive Matrices Test (1962 Revision) and Otis Quick-Scoring Mental Ability Gamma Test, Form C.

RPMT				Otis							
Score	N	Score	N	Score	N	Score	N	Score	N	Score	N
14	1	25	2	43	1	54	1	65	1		
15	1	26	4	44	1	55	2	66	4		
16	—	27	4	45	—	56	1	67	1		
17	1	28	4	46	—	57	1	68	2		
18	1	29	1	47	1	58	3	69	3		
19	3	30	5	48	2	59	2	70	1		
20	2	31	2	49	1	60	1	71	1		
21	—	32	1	50	1	61	—	72	—		
22	4	33	—	51	—	62	2	73	1		
23	3	34	—	52	1	63	1	74	2		
24	5	35	1	53	2	64	1	75	1		
								76	3		

ance. This was a considerably greater range of scores than that obtained from the RPMT.

The correlation between the RPMT and the Otis was .27 ($.05<p<.10$).

Code-Lock Performance and Intelligence. The intercorrelational matrix of the five code-lock measures and the two intelligence measures is shown in Table 2. These correlations are based on averages across all difficulty levels for the code-lock measures.

Only one of the five correlations between the code-lock measures and the RPMT was significant at the .05 level of confidence. This was the measure of time required to solve the problem after having found the first correct button in a problem sequence ($r=-.31$). The correlation between the actual number of errors minus the expected number of errors approached, but did not achieve, significance, ($r=-.29$; $.05<p<.10$).

The correlations for each RPMT measure with each of the three levels of expected errors are shown in Table 3 for the RPMT. The smaller predicted number of errors led to higher correlations. Four of the five correlations were significant at the .05 level of confidence for the expected error range of 1 to 4 errors; the fifth correlation approached significance ($r=-.29$; $.05<p<.10$). None of the correlations for either the 5 to 6 or the 7 to 9 error ranges was significant.

Three of the five correlations between the Otis and the code-lock measures were significant. These code-lock measures were: time per response

TABLE 2.—Correlation matrix for intelligence and code-lock measures.

	Measure[a]					
	Raven	M_1	M_2	M_3	M_4	M_5
Otis	.274	−.613**	−.196	−.245	−.594**	−.574**
Raven		−.246	−.293	−.182	−.178	−.311*
M^1			.204	.246	.832**	.962**
M^2				.597**	.136	.373**
M^3					.424**	.276
M^4						.705**

[a] M^1—Mean time per response (1st and 2nd solutions combined)
M^2—Mean surplus errors (1st solution)
M^3—Proportion of errors (1st and 2nd solutions combined)
M^4—Mean time per response (2nd solution)
M^5—Mean solution time after 1st correct response (1st solution)
* $=p<.05$ ** $=p<.01$

TABLE 3.—Correlations between code-lock measures for three levels of expected errors (problem difficulty) and both the Raven Progressive Matrices Test and the Otis Quick-Scoring Mental Ability Gamma Test, Form C.

Test	Expected Errors	Measure[a] M_1	M_2	M_3	M_4	M_5
	1-4	−.332*	−.354*	−.298	−.304*	−.397**
Raven	5-6	−.154	−.058	−.085	−.059	−.270
	7-9	−.260	−.023	−.048	−.172	−.268
	1-4	−.565**	−.106	−.168	−.500**	−.488**
Otis	5-6	−.586**	−.132	−.185	−.494**	−.551**
	7-9	−.590**	−.117	−.010	−.516**	−.522**

[a] See Table 2 for description of measures
*=$p<.05$ **=$p<.01$

for first and second solutions combined (r=.61), time per response for the second solution only (r=−.59), and solution time after finding the first correct button in the sequence (r=−.57). The intercorrelations among the three code-lock measures that correlated significantly with the Otis were quite high, ranging from +.70 to +.96.

The correlations with the Otis for each measure at each of the three levels of expected errors are shown in Table 3. The relative magnitudes of the correlations were not appreciably affected by the breakout of different expected numbers of errors. However, the correlations were slightly lower than for the data based on all problems, perhaps because of the decreased stability afforded by the smaller number of problems per data point for each subject for a given expected error range.

IV. Discussion.

The small, non-significant correlations (.27) between the Otis and the RPMT is consistent with previous research findings.[e.g., 6] On the other hand, it is probable that such correlations indicate that the two tests do not generally measure the same aspects of mental abilities. Also, in this study the range of scores on the RPMT was rather narrow, which may have further limited the RPMT-Otis correlations as well as the RPMT and code-lock correlations. The finding that the RPMT scores were uniformly at a high level was not expected since the 1962 revision of the RPMT was developed to provide a non-verbal assessment of ability in adults of above-average intellectual ability.[8] Instead, the present findings suggest that the 1962 version of the RPMT also has a relatively low functional "top," at least when applied to college student populations and may, therefore, not be appropriate for use with individuals of above average intelligence.

In general the code-lock measures did not relate closely to RPMT performance, except at the least difficult problem level. However, even with the least difficult problems the correlations, although significant for the three time measures and one error measure, did not exceed .40. In comparison, Alluisi and Coates[1] found correlations ranging to .52. As noted above, the general lack of a strong relationship in this study may have been due in large part to the limited range of RPMT scores. In the Alluisi and Coates experiment, the range of RPMT scores was considerably greater, as according to Coates,[4] only 40% of their subjects exceeded the 90th percentile, while in this study 80% exceeded that level of performance. It is particularly puzzling, however, as to why the significant correlations were obtained only for the least difficult problems. Perhaps the differentiation which did occur in RPMT scores was related in some manner to the subjects' ability to estimate an entire sequence for the least difficult problems from a few button presses, while this was not possible for any subjects with the more complex response sequences.

In comparing this study with that of Alluisi and Coates,[1] the procedural variations between the code-lock and COTRAN tasks should also be noted. First, the instructions given COTRAN

subjects did not mention speed or accuracy during the two phases consisting of code-lock problems. In this study, each subject was told to respond as rapidly and as accurately as he could. Second, the response keys (not buttons) for the COTRAN task were arranged for convenient operation by the fingers and thumb of the right hand, while subjects used only the right hand index finger in this experiment. The emphasis on speed and accuracy in this setting may have increased the intellectual demand upon the subject considerably from that in the COTRAN task, and the difference in response devices may have increased the influence of manual-dexterity in the COTRAN findings. How these factors relate to the relative performance on the intelligence measures is not clear.

With respect to the Otis, it was found that the three time measures were reasonably efficient predictors of Otis performance at all three difficulty levels. It is noteworthy that the magnitude of these correlations approached those usually obtained when the Otis is compared with other measures of intelligence and achievement.[3] It is not clear why the Otis should correlate only with time, and not error, measures. Perhaps the fact that the Otis is administered under a time limit is important to the relationship. The RPMT, which was also administered as a speed, rather than power, device showed some similar trends where significant correlations occurred, and thus is consistent with the Otis findings in indicating the significance of the time factor as a basis for the obtained relationships. In any event, the code-lock task does seem related to the Otis as a measure of general intelligence, and therefore may offer some promise as an assessment device which does not require highly developed verbal skills. If so, it would have potential utility in the assessment of ability to learn where orthodox evaluation techniques discriminate against individuals whose experiences have not encouraged the development of the specific language facilities needed for successful test performance.

REFERENCES

1. Alluisi, E. A. and G. D. Coates: A code transformation task that provides performance measures of non-verbal mediation (COTRAN). National Aeronautics and Space Administration Contractor Report No. 895, 1967.
2. Chiles, W. D., A. E. Alluisi, and O. S. Adams: Work schedules and performance during confinement, HUMAN FACTORS, 10:143–196, 1968.
3. Chronbach, L. J.: *Essentials of Psychological Testing, 2nd Ed.*, New York, Harper and Brothers, 1960, 215–225.
4. Coates, G. D.: Personal communication, 1970.
5. Estes, B. W.: Relationships between Otis, 1960 S-B, and WISC, JOURNAL OF CLINICAL PSYCHOLOGY, 21: 296–297, 1965.
6. Estes, B. W., M. E. Curtin, R. A. DeBurger, and C. Denny: Relationships between 1960 S-B, 1937 S-B, WISC, Raven,, and Draw-A-Man, JOURNAL OF CONSULTING PSYCHOLOGY, 25: 388–391, 1961.
7. Gaw, F.: A study of performance tests, BRITISH JOURNAL OF PSYCHOLOGY, 15: 374–392, 1925.
8. Raven, J. C.: Advanced Progressive Matrices, Sets I and II. Plan and use of the scale. London, H. K. Lewis & Co. Ltd., 1965.

APPENDIX I

Intelligence Measure Description

Otis Quick-Scoring Mental Ability Gamma Test, Form C

This test is a revision of one of the first devices to be used for the assessment of the intelligence levels of individuals examined in groups. The particular form employed is suitable for use with persons who have at least reached high school age. The 80 items sample such areas as vocabulary, quantitative reasoning, perception of spatial relationships, and abstract reasoning. The test is not so sophisticated as some more recently developed intelligence assessment techniques; however, its predictive validity for academic achievement has been shown to compare favorably with most other tests. It is easy to both administer and score, and is especially useful in situations where time for assessment is limited.

Ravens Progressive Matrices Test (1962 Revision)

The RPMT was developed as a device to measure intelligence independently of an individual's level of verbal skill. Although widely used in England, this technique has had limited application in the United States because it has lacked suitable norms and has not been found to correlate highly with academic performance. It has been used primarily as a means of identifying individuals with good reasoning ability who have poor reading or language development. The items employed in the RPMT are essentially two-dimensional analogy problems. For each problem, a matrix of eight patterns in a three-row by three-column arrangement is presented the subject. The subject then attempts to choose the correct pattern from several alternatives to complete the ninth position in the matrix. This requires the determination of two principles, one which governs the transition in patterns across rows, and one which determines the variation down columns. The two principles can then be applied to the available alternatives to determine the correct choice for the final pattern. The version of the RPMT used in this experiment is the Advanced Set which is suitable for use with adults.